(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification⁷: C12N 15/80

(21) International Application Number:
PCT/GB2003/004716

(22) International Filing Date: 31 October 2003 (31.10.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0225390.4   31 October 2002 (31.10.2002)   GB

(71) Applicant (for all designated States except US): HOR-TICULTURE RESEARCH INTERNATIONAL [/GB]; Wellesbourne, Warwick, Warwickshire CV35 9EF (GB).

(72) Inventors; and
(75) Inventors/Applicants (for US only): BURTON, Kerry [/GB]; Horticulture Research International, Wellesbourne, Warwick, Warwickshire, CV35 9EF (GB). CHALLEN, Michael [/GB]; Horticulture Research International, Wellesbourne, Warwick, Warwickshire, CV35 9EF (GB). ELLIOTT, Timothy [/GB]; Horticulture Research International, Wellesbourne, Warwick, Warwickshire, CV35 9EF (GB). SREENIVASAPRASAD, Surapareddy [/GB]; Horticulture Research International, Wellesbourne, Warwick, Warwickshire, CV35 9EF (GB). EASTWOOD, Daniel [/GB]; Horticulture Research International, Wellesbourne, Warwick, Warwickshire, CV35 9EF (GB).

MALLOY, Shannon [/GB]; Horticulture Research International, Wellesbourne, Warwick, Warwickshire, CV35 9EF (GB).

(74) Agent: LORD, Hilton, David; Marks & Clerk, 57-60 Lincoln's Inn Fields, London WC2A 3LS (GB).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: SELECTIVE EXPRESSION IN FILAMENTOUS FUNGI

(57) Abstract: Gene promoters activated only during development of the fruiting body of filamentous fungi permit expression of large quantities of heterologous genes, even where such are otherwise toxic to the organism.

SELECTIVE EXPRESSION IN FILAMENTOUS FUNGI

The present invention relates to the transformation of filamentous fungi with heterologous DNA, methods for such transformation and processes for the collection of the resulting expression products.

Filamentous fungi have an unusual life cycle, in that the majority of the cycle is spent as mycelium, which is a largely hidden network of filaments permeating throughout and over a suitable food source. In the case of *Agaricus bisporus*, or button mushroom, which is a horticultural adaptation of a wild mushroom, the food source is decomposed leaf litter, in the wild, or generally composted cereal straw, when cultivated, and little signs of growth can be observed above ground level, other than the characteristic smell, until the fruiting bodies, or sporophores, start developing.

Compared with the culture of other vegetables, the commercial production of button mushrooms is complex. They are best grown indoors in controlled environmental conditions in trays, shelves or bags filled with a speciality compost, specifically made by, and for the, the mushroom industry from ingredients including: plant remains, such as cereal straw, corn husks, hay, rape straw; animal manure, such as from horse, chicken, cow, pig; gypsum, and other additives.

Mushroom compost is made by the mixing of the raw ingredients, for example straw, manure and gypsum base, with water, which leads to microbial degradation and heat production. The composting process has two phases. In Phase 1 composting, the mixture is either stacked in a windrow outdoors, or in a barn, and turned every 2 to 3 days over a three-week period, or it is placed in a walled bunker and aerated from below over an 8-10 day period. This compost is then subjected to Phase II composting in an aerated bunker where the temperature is raised to 58-60°C for approximately 12 hours and held at 50-54°C for 5 days. The compost is then ready for spawning.

Spawn is, typically, sterilised cereal grain, such as rye or millet, which has been colonised with the mushroom mycelium. Spawn is mixed with the compost and

maintained at 24°C. After approximately 17 days at this temperature, the compost is fully colonised by the mycelium.

A casing layer is then applied to the surface of the colonised compost to a depth of approximately 5 cm. Casing normally consists of peat (or peat substitute for example, coir, bark, mineral washing), chalk/lime/sugarbeetlime, water and a small amount of specialised spawn, which may be the mycelium cultured on compost, grain, vermiculite or other substrate. When the surface of the casing layer is colonised with mycelium, the temperature is lowered to 16-18°C, which induces production of mushroom sporophores, or fruiting bodies. From application of the casing to harvesting the first sporophore takes approximately 17 days.

The production of mushrooms occurs periodically, over cycles known as flushes, at 7-10 day intervals. High productivity farms normally crop mushrooms over 3 flushes. The time from the first harvest of the first flush to the last harvest of the third flush is about 21 days.

The time each growing house is occupied with each crop is from after casing and during crop production, or 17 + 21 = 38 days, or thereabouts. This allows a growing house to be used for 9 crops per year. High productivity farms routinely have yields of 25kg/m². Because mushroom production is not light dependent, the crops are grown in layers, either four or six to maximise productivity. Accordingly, the four-layer system is expected to produce 3,200 tonnes per annum in a hectare of growing rooms.

The rapid development of the fruiting body is unique in the vegetable kingdom, and represents a significant target for the expression of suitable genes. If this rapid, volumetric expansion of the fruiting body could be harnessed together with expression of a suitable gene, then this might make possible the production, on an industrial scale, of biological products that currently defy large scale production.

Attempts to express heterologous proteins in filamentous fungi have met with numerous problems, not least being the difficulty associated with transforming the fungi in the first place. For example, WO95/02691 relates to transforming a mushroom mycelium and fruit bodies through methods such as electroporation with suitable vectors. Although this method works on a small scale, it is not especially efficient.

3

WO98/45455 relates to the possibility of transforming moulds, such as *Agaricus bisporus*, with the *Agrobacterium tumefaciens* bacterium, which causes crown gall tumours at the wound site of infected dicotyledonous plants. This bacterium is well known for its ability to transform plants, but it has only recently been established that it can also transform filamentous fungi.

WO 96/41882 discloses the expression of hydrophobins by the *hyp A, hyp B, hyp C* and *hyp D* genes. These are naturally occurring fungal products, and are expressed during fruiting. Heterologous expression is suggested, in connection with enhancing the flavour and/or nutritional content of the fruiting bodies.

Mol. Gen. Genet. (1993), *238*: 91-96, provides a reporter-gene system in *S. commune*, for the expression of hydrophobins.

The problem still remains that the expression of heterologous genes in substantial amounts will generally substantially reduce the growth potential of the mycelium and, therefore, the harvest of the fruiting body, and prohibits the expression of any substance which is, in any way, toxic to the growth of the fungus.

Surprisingly, it has now been found that at least three genes are switched on, or otherwise subjected to elevated levels of expression, at around the veil-break stage of fruiting body development, and that heterologous DNA under the control of the expression mechanisms of these fungal genes can be selectively expressed at this stage of development of the fungus, rather than during growth of the mycelium.

Thus, in a first aspect, the present invention provides a filamentous fungus transformed with a heterologous sequence of DNA, the fungus being capable of expressing the heterologous DNA, characterised in that the heterologous DNA is under the control of a filamentous fungus transcription promoter active substantially only during stage 1, or later, of the development of the fruiting body of the fungus.

It is a particular advantage of the present invention that little or no metabolic energy need be diverted from mycelium growth, thereby maximising fruiting body mass and concomitant tissue capable of expressing the heterologous gene once it is switched on.

4

It is a further advantage of the present invention that little or none of the heterologous gene product is expressed during vegetative growth of the mycelium, thus enabling the production of substantially any substance capable of expression in the filamentous fungus in question, even if that substance, either alone or in combination, results in the death or stasis of the fungus. The promoters allow the synchronous switching on of the gene at a time of rapid growth and high metabolism so that, by the time any potentially toxic effects become apparent, harvestable quantities of the substance are available. Where the gene product is a regulator of mushroom growth, for example, then such considerations are not generally necessary.

It is particularly preferred to harvest the expression product of the heterologous DNA, and the product may be purified, if and as desired, by any suitable means, such as are well known in the art.

Accordingly, the present invention further provides a method for the production of a substance expressible by a DNA sequence, wherein the sequence is operably associated with a filamentous fungus transcription promoter active substantially only during stage 1, or later, of the development of the fruiting body of the fungus, the sequence and promoter being expressibly incorporated in a filamentous fungus, the fungus being cultured to fruition and the product being harvested.

At present, no less than three genes have been identified that are expressed substantially only during development of the fruiting body, and particularly during stages 4 to 7 (veil-break onwards). Without being constrained by theory, it is likely that these genes are associated with the massive water uptake required for the expansion of the fruiting body and its maturation.

The three genes, so far identified, are *abst1*, *rafe* and *mag2*. The expression product of *abst1* appears to be involved in the transport of sugars, whilst the expression product of *rafe* is a putative riboflavin aldehyde forming enzyme. The expression product of *mag2* is a so far unidentified morphogenesis associated protein.

The sequences of the three genes, and associated promoter and terminator regions, are given in the accompanying Sequence Listing.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 illustrates the stages of development of the fruiting body of a mushroom fungus, from vegetative mycelium (VM) to Stage 1 through to Stage 7 .

Figure 2A illustrates the level of expression of *abst1* during the various stages of development of the fruiting body, from vegetative mycelium (VM) to Stage 1 through to Stage 7.   Figure 2B illustrates the gene expression associated with the various parts of the fruiting body at stage 4. US represents upper stipe; LS represents lower stipe; C represents cap; G represents gill; and S represents skin.

Figure 3 corresponds to Figure 2, except showing the expression of *rafe*. The stages in Figure 3A are as in Figures 1 and 2A, above. In Figure 3B, the parts of the fruiting body in which expression of *rafe* are shown, is as in Figure 2B, above.

Figure 4 corresponds to Figures 2 and 3 above, showing the expression of *mag2*. The stages and fruiting body parts showing expression are as above.

Figure 5 shows a construct with restriction enzyme sites represented by numbers 1-6.

Figure 6 shows a promoter - intron cassette, comprising *A. bisporus GPD* promoter 5'UTR sequence, an ATG start codon and three naturally occurring introns (*in1*, *in2* and *in3*). The restriction sites *Kpn*I and *Nar*I are shown. The nucleotide sequence of this cassette is given in SEQ ID NO. 12. The ATG (start codon) and VKV.....()F().......GRIG(R) have not been included in the complete amino acid sequence.

Figure 7 shows a leader (*L*) sequence, the *SPR* leader sequence, comprising 57 nucleotides (SEQ ID NO. 13) encoding 19 amino acids (SEQ ID NO. 14). This leader sequence is derived from Y13805 deposited as Accession Number Y13805 in the EMBL/GenBank/DDBJ databases (Kingsnorth, C. S., Eastwood, D. C. & Burton, K. S. (2000). *Agaricus bisporus* partial mRNA for serine proteinase).

6

Figure 8 shows a Terminator cassette, corresponding to Figure SEQ ID NO. 35, comprising a *A. nidulans trpC* terminator 3'UTR and a number of engineered restriction sites (*Bam HI, BglII* and *KpnI*).

Figure 9 shows a construct in a pBluescript plasmid, combining *A. bisporus GPD* promoter with introns (as detailed above, referred to as *P*) fused with *eGFP* gene and *A. nidulans trpC* terminator (as described above, referred to as *T*). Engineered *KpnI* sites are used to excise expression cassette from pBluescript and transfer to *Agrobacterium* binary vector harbouring hygromycin (*hph*) resistance selectable marker. T-DNA introduced into *A. bisporus* by Agro-transfection contains *hph* and heterologous protein (GFP) as a pair of divergently transcribed genes.

It will be appreciated that the Basidiomycetes, including members of the *Agaricales*, of which *A. bisporus* is one, share the exceedingly rapid development of the fruiting body in common. Without being constrained by theory, it is envisaged that this development is as a result of a rapid increase in osmotic pressure in the cells of the immature fruiting body, thereby causing a rapid influx of water into the cells. The resulting sudden expansion of the cells expands the fruiting body up to several hundred times its original size.

One or more sugar transport mechanisms are switched on at the early stages of fruiting body development, and abundant expression product is noted, especially by stage 4. These genes form a preferred subject of the present invention, and especially the control element associated therewith, but it will be appreciated that any gene selectively expressed, or with greatly enhanced expression, during development of the fruiting body is useful in the present invention.

The *abst1* gene is up-regulated, by more than 100-fold, during mushroom development, and is abundantly expressed through stages 4 to 7, and represents about 0.6% of the transcripts detected at stage 4. The transcript is about 1.7kb in length.

7

The gene product of *rafe* is up-regulated by about 50-fold during mushroom development, and is abundantly expressed from stage 4 onwards, as with *abst1*. The transcript is 0.7kb in length.

The expression product of *mag2* is up-regulated by about 30-fold during mushroom development, and represents about 0.2% of the transcripts at stage 4. Unlike *abst1* and *rafe*, the expression of *mag2* appears to be comparable in both the stipe and cap tissues. The length of the transcript is about 0.7kb.

The control elements, and especially the promoters, of these and other genes expressed during the development of the fruiting body are particularly useful in the present invention. Elements from genes associated with sugar transport are particularly preferred.

It is a particular advantage of the present invention that heterologous genes can be expressed at selected stages of sporophore development, where these genes might otherwise be harmful to the fungus.. Expression of the heterologous genes occurs substantially only during growth of the fruiting body so that, unless the gene product is acutely toxic, then large amounts of the gene product can be expressed which would otherwise harm or hinder growth of the fungus.

Owing to the commonality of the fruiting body growth process, the promoters of any one filamentous fungus, switched on during the fruiting body growth cycle, may be employed in other filamentous fungi in the context of the present invention, in order to express heterologous genes.

In the accompanying Sequence Listing, SEQ ID NO's 12 and 13 are the promoter sequences associated with *abst* 1 and *rafe*, respectively, and the full sequences are provided for *abst1* and *rafe* as SEQ ID NO's 4 and 9. These include a substantial portion of the upstream promoter area and, following the procedures described in the accompanying Example, the upstream promoter area is also obtainable for *mag2*.

The promoter may be used in association with other suitable control sequences, such as terminators. A suitable terminator may be as shown in the accompanying

8

sequences, or may be the Aspergillus nidulans trpC terminator, for example. Other terminators are well known in the art.

The terminator sequences of *abst 1* and *rafe* are provided as SEQ ID NO's 35 and 36, and it has been found that it is desirable to use terminators generally associated with the promoters, so that it is preferred to use the *abst 1* terminator sequence with the *abst 1* promoter sequence, and likewise for *rafe*. However, it will be appreciated that the present invention also envisages any suitable expression system comprising the promoter, and any suitable terminator may be employed, as desired.

It will also be appreciated that the promoter sequences and terminator sequences of the invention are preferably those as listed as SEQ ID NO's 12 and 13, and 35 and 36, respectively, and sequences comprising these sequences, as well as sequences hybridising with these sequences, preferably under conditions of 60% stringency or higher, provided that promoter activity is retained in the by the sequence or the sequence to which it hybridises. Mutations and naturally occurring variants of the sequences are encompassed, and it may be, for example, appropriate to introduce a restriction site or sites for ease of manipulation. Provided that promoter activity is retained, there is no restriction on how much the promoter may be modified. Similar considerations apply to the terminator sequences.

The promoter regions may be used in their entirety when preparing heterologous genes for expression in filamentous fungi. Alternatively, it may be preferred to use consensus sequences from these regions. There is no especial advantage to using consensus sequences, except that these may be shorter. Otherwise, it is sufficient to supply the promoter upstream of the desired heterologous gene. Being a promoter, there is also no requirement that it be in the correct reading frame, just within the appropriate promoter distance.

It is generally preferable that the transformed fungus also expresses a linked selectable marker. Any marker known in the art may be used, and may be excised once a faithful strain has been generated. However, it is generally preferable that the transformed fungus maintains a marker to ensure that the desired heterologous product is still produced, and to ensure that there is no reversion to wild type. In this respect, it

is preferred that the marker have no significant negative effect on either the fungus or the product. Such markers may normally be selected from resistance markers, in order that the growth medium contain amounts of an antifungal agent ensuring that only transformed fungus can grow successfully.

Suitable markers include the hygromycin resistance cassette and the benomyl resistance tubulin gene.

Suitable methods for transforming filamentous fungi are as described above with respect to WO95/02691 and WO98/45455, which disclosures are incorporated herein by reference.

In general, the desired control sequences are ligated with the appropriate heterologous expression sequences and prepared for insertion into a suitable preparation of the fungus, such as protoplasts, all by methods well known in the art.

The resulting organism can then be grown by standard methods, and prepared as spawn after cultivation of the resulting mycelium. Spawn has the advantage that it can be stored inert for relatively long periods of up to about a year, although it is generally preferred to use it within about 4 months.

Spawn may be produced in any recognised manner, such as by growing the mycelium on sterile agar and introducing the culture to autoclaved grain. The grains may then be stored at elevated temperature to encourage colonisation, and then kept at reduced temperatures until needed.

It will be appreciated that the heterologous gene for incorporation may be in the form of cDNA or genomic DNA. cDNA is preferred, as it is generally shorter and more easy to handle.

It will also be appreciated that the heterologous gene insert should encode the sequence desired, including leader sequences and cleavage sequences, if required.

It will also be appreciated that greater expression may be achieved if fungal codons are used in place of mammalian codons, although expression will still occur, and such substitution is not necessary.

10

It will further be appreciated that heterologous genes may need to be expressed in the form of a cassette, for example, in order to produce the required product. In general, it is preferred to require as few heterologous gene products as possible, as the greater the number, the more likely it is that the fungal metabolism will interfere in some way, and it is generally desirable to minimise unpredictability.

Thus, although not limited thereto, it is generally preferred to limit the number of heterologous expression products to one, two or three, preferably one or two, and preferably one, other than any marker. The marker is preferably linked to the heterologous gene, such as downstream of the gene and also under the control of the fungal promoter, so as best to indicate successful and/or continuing stable transformation.

It is particularly preferred that products such as peptides be the target, as these can be harvested relatively simply. Thus, enzymes and antibodies are particularly useful, although conformational proteins, such as vaccine antigens, and active peptides such as interferons are also useful.

Accordingly, heterologous genes suitable for expression in the filamentous fungi include those whose expression results in the production of: antibodies, including other diagnostic material; secondary metabolites, such as lectins, pesticidal compounds such as *Bacillus thuringiensis* toxin (Bt toxin); therapeutic compounds such as vaccines, steroids, heterocyclic organic compounds; biological macromolecules, such as interferon, endostatin and insulin; and medical enzymes, such as thrombolytics and cerebrosidases.

In the context of the present invention, the term "heterologous" includes native DNA not normally associated with heightened expression during sporophore production. In such a respect, the native gene becomes heterologous insofar as its expression pattern is altered. Such expression may generally serve one of two purposes. The first is generally to obtain large/greater amounts of native protein, such as by transforming the filamentous fungi with extra copies or modified copies of a native gene or genes. The second may be used instead to affect/control the characteristics of mushroom crop production, such as by altering the timing of crop, flushing pattern,

11

yield, growth rate and/or final size of the mushroom sporophore. This latter may also suitably be achieved by the introduction of heterologous DNA from other species, if desired.

The crops are preferably allowed to go to full cap development, where possible, in order to maximise expression of the heterologous gene, although the skilled person will appreciate the best stage for harvesting any given product. The resulting caps may then be processed in any suitable manner to extract and/or purify the product, or the caps may otherwise be employed or processed, as desired.

Where the product is potentially dangerous, standard procedures may be employed between crops to entirely sterilise the area, such as steam sterilisation and swabbing of the walls, as described above.

The present invention will now be illustrated further, by reference to the following, non-limiting Examples.

EXAMPLE 1

Mushroom strains and growing conditions

A commercial *A. bisporus* strain U3 (Sylvan, U.K.) and a carboxin resistant *A. bisporus* mutant C54-*carb*.8 were used in this work. Vegetative mycelium was produced on sterile compost at 25°C and fully colonised compost (21 days) was frozen in liquid nitrogen. Mushrooms were grown in trays according to commercial practice at the Horticulture Research International mushroom cropping unit. Sporophores were produced in synchronous weekly flushes and mushroom fruit bodies from second flush were harvested at developmental stages one to seven and flash frozen. For tissue expression analysis, stage 4 mushrooms were dissected into stipe (upper and lower), cap (pileus trama), skin (pilei pellis) and gills (lamellae) and frozen. All frozen samples were stored at −80°C.

Bacterial strains, vectors and phagemid/cosmid DNA extraction

Escherichia coli strains XL-1 Blue and XLOLR (Stratagene) were used for the preparation and propagation of cDNA clones. Phagemid and Cosmid DNA extractions were carried out using the Tip 20 plasmid DNA extraction kit (Qiagen).

Preparation of an ordered library of random cDNA clones

Total RNA extraction and poly (A)$^+$ RNA isolation were carried out as described previously. Using 5 μg poly (A)$^+$ RNA, cDNA libraries (ZAP EXPRESS, Stratagene) were constructed from veil break stage mushrooms, and mushrooms were harvested and stored for two days. Mass excision of the cDNA libraries was performed according to the manufacturer's (Stratagene) instructions. Clear single colonies (cDNA clones in pBK-CMV) were picked and ordered in microtitre plates containing 200 μl media 96 broth with Kanamycin (50 μg ml^{-1}) in each well.

Differential screening

Duplicate nylon membranes containing DNA from the clones were prepared for differential screening. For preparing probes cDNA was generated from 5 μg poly (A)$^+$ RNA using the Ready-to-Go T-primed cDNA synthesis kit (Amersham Pharmacia Biotech). Total cDNA's were labelled with [α-^{32}P] dCTP using the Redi-prime random labelling kit (Amersham Pharmacia Biotech). One set each of the membranes containing 3500 cDNA clones from the veil break stage mushrooms were hybridised with the cDNA probe from the veil break stage mushrooms and the cDNA probe from button stage mushrooms. Putative differentially expressed cDNA clones were re-screened to reduce false positives.

Northern analysis

Total RNA was extracted from compost colonised vegetative mycelium and seven different developmental stages of the mushroom, freshly harvested button stage mushroom and samples stored from 1 – 5 days as well as from the stipe, cap and gill tissue. Equal quantities (15 µg) of total RNA for each sample were electrophoresed and Northern blots prepared as described elsewhere. cDNA randomly labelled with [α-^{32}P] dCTP, as described above, was used as the probe.

Genomic library screening and identification of the gene

For isolating cosmid clones containing the gene, a genomic library constructed from a carboxin resistant mutant of *A. bisporus* C54-*carb*.8 in cosmid vector Lawrist was used. Preliminary screening of DNA pools from 56 microtitre plates (96 clones each, *ca*. 5376) was done by PCR using primers designed from the cDNA sequence. Individual cosmid clones containing the gene were identified by probing colony blots of the 96 clones in each of the positive pools.

Nucleotide sequence determination and analysis

cDNA's in the phagemid vector pBK-CMV were initially sequenced using the vector primers T3 and T7 and the full double strand sequence was obtained using additional primers synthesised from known sequence. Genomic sequence was generated from cosmid clones by primer walking, where the initial sequencing was by done using the cDNA primers and further primers for sequencing were designed from known sequence. Nucleotide sequences were determined using the ABI automated DNA sequencing technology. Sequencing reactions were carried by thermal cycling using the ABI Prism™ BigDye terminator cycle sequencing kit (ABI-Perkin Elmer) as per manufacturers' instructions. Editing and assembling of the sequence data were done using the programmes within the DNASTAR package (Lasergene software, Dnastar

14

Inc.). Homology searches and nucleotide and amino acid sequence comparisons were made using a suite of software available on the WWW, particularly ExPASy (www.expasy.com) and other linked sites.

EXAMPLE 2

Constructs

Constructs are prepared to maximise expression of specific proteins through the use of homologous regulatory sequences from the mushroom *Agaricus bisporus*. The sequences comprise (see Figure 5):

- *promoter sequences P*: 5' untranslated regions preceding *A. bisporus* genes previously shown to be highly expressed in specific mushroom tissues or developmental stages;

- *leader sequences L* (optional): specific N-terminal regions of proteins that encompass the ATG nucleotide start codon for the Methionine amino acid and a defined number of other amino acids implemented in the cellular 'targeting' of proteins;

- *intron sequences I* (optional): the protein coding sequence is interrupted with a short region(s) of non-coding sequence. Intron sequences are naturally occurring in many eukaryotes and have been defined in all *A. bisporus* genes hitherto characterised. In the absence of *L* sequences the *I* cassette encompasses an ATG start codon that precedes the intron sequences;

- *gene/protein sequences G*: coding sequences for the desired heterologous protein and encompassing a stop codon;

- *terminator region T*: 3' untranslated regions following defined genes from *A. bisporus*, or other fungi, and implemented in the accurate termination and stability of gene transcripts.

P, L I, G and *T* sequences are engineered from defined nucleic acids using the 'polymerase chain reaction' with defined oligonucleotide primers to introduce specific and novel restriction enzyme sites that facilitate the assembly of constructs through

15

molecular cloning protocols. An example construct is shown in Figure 5 with restriction enzyme sites represented by number 1-6. Therefore, cassette *P* with engineered restriction sites 1 at the 5' end and 2 at the 3' end is ligated with cassette *L* with engineered sites 2 (or compatible overhang sequence) at the 5' end and 3 at the 3' end.

P, *(L)*, *I*, *G* and *T* cassettes are assembled in a plasmid vector (for example pBluescript2) and amplified at appropriate stages in *E. coli* using appropriate molecular cloning protocols.

Promoter (*P*) sequences used in heterologous protein constructs:

The *ABST1* promoter sequence is shown in SEQ ID NO. 12. The *RAFE* promoter sequence is shown in SEQ ID NO. 13.

Strategy for development of promoter - intron cassette

A. bisporus GPD promoter 5'UTR sequences ATG start codon and three naturally occurring introns (in1-3; lower case). An Engineered PCR product (486 nucleotides) with restriction sites *Kpn*I and *Nar*I and introns is shown in SEQ ID NO. 14 and in Figure 6.

Leader (*L*) sequence

An example of a suitable leader sequence is the *SPR* leader sequence, comprising 57 nucleotides (SEQ ID NO. 15) encoding 19 amino acids (SEQ ID NO. 16), and deposited as Genbank (NCBI)Accession Number Y13805. The sequence alignment of the nucleotide and amino acids is given in Figure 7.

Intron (*I*) sequences:

In addition to *GPD* introns referred to above, numerous *A. bisporus* introns have been characterised that can be used as intron cassettes (see tables 2 to 5, below). These introns have been established by comparison of partial cDNA and genomic DNA

16

sequences. The introns are described by their gene of origin (i.e. putative gene product) and Genbank (NCBI) cDNA accession number. For each gene, introns are numbered from the 5' end. The first identified introns of each gene are theoretical and unconfirmed (cDNA sequence not available at 5' end and therefore identified by comparison of gene sequence with sequences of similar genes from other organisms) and are not included in this list.

Terminator (T) sequences:

The *abst1* terminator sequence is given in SEQ ID NO. 35. The *rafe* terminator sequence is given in SEQ ID NO. 36.

Strategy for development of terminator (T) cassette

A. nidulans trpC terminator 3'UTR as an engineered PCR product (comprising 800 nucleotides) with introduced restriction sites *BgI*II and *Kpn*I, is given in SEQ ID NO. 37, and shown diagrammatically in Figure 8.

Construct strategy for expression of GFP protein:

Construct in pBluescript plasmid combines *A. bisporus GPD* promoter with introns (as detailed above; *P*) fused with *eGFP* gene and *A. nidulans trpC* terminator (as described above; *T*), and represented in Figure 9. Engineered *Kpn*I sites are used to excise expression cassette from pBluescript and transfer to *Agrobacterium* binary vector harbouring hygromycin (*hph*) resistance selectable marker. T-DNA introduced into *A. bisporus* by Agro-transfection contains *hph* and heterologous protein (GFP) as a pair of divergently transcribed genes.

In the accompanying Sequence Listing, it will be noted as follows:

Organism:	*Agaricus bisporus*
Gene designations (tentative):	*abst1* (sugar transporter); *rafe* (putative riboflavin aldehyde forming enzyme gene); *mag2* (unidentified morphogenesis associated gene)
Gene identification:	By differential screening of a cDNA library (3500 random clones) from stage 4 mushrooms
Genomic clones:	From a genomic library of *A. bisporus* strain C54-*carb*8 in cosmid vector Lawrist (*ca.* 30 – 40 kb genomic fragments)
abst1:	Up-regulated (more than 100 fold) during mushroom development, abundantly expressed through stages 4 – 7 (later stages of mushroom development), *ca.* 0.6% transcripts at stage 4, represented by 20 clones in the differential library, 1.7 kb transcript
rafe:	Up-regulated (up to 50 fold) during mushroom development, abundantly expressed through stages 4 – 7, 0.7 kb transcript
mag2:	Up-regulated (up to 30 fold) during mushroom development, *ca.* 0.6% transcripts at stage 4, represented by seven clones in the differential library, comparable levels of expression in stipe and cap tissue, ca. 0.7 kb transcript

The following Table provides an index of the accompanying SEQ ID NO's, and gives an indication as to their identity. The sequence provided in SEQ ID NO 4 was obtained on the basis of single strand analysis and is for guidance only, as there are some errors and unknown nucleotides. The promoter sequence of SEQ ID NO. 12 is preferred, as it is shorter, and does not comprise as many restriction sites. In addition, it was sequenced as a double strand.

18

Table 1

Sequence description	SEQ ID NO
Abst1 cDNA sequence	1
Abst1 protein encoded by ORF of SEQ ID NO 1	2
abst1 Genomic sequence	3
abst1 Promoter region (approx. 1000 bases upstream of ORF).	4
abst1 Terminator region (approx. 700 bases downstream of ORF	5
rafe cDNA sequence	6
Rafe protein encoded by ORF of SEQ ID NO2	7
rafe Genomic Sequence	8
rafe Promoter region (approx. 1000 bases upstream of ORF)	9
rafe Terminator region (approx. 1200 bases downstream of ORF):	10
. *mag2*: Partial cDNA sequence	11
abst 1 promoter sequence	12
Rafe promoter sequence	13
Engineered PCR product (486 nt) with restriction sites *Kpn*I and *Nar*I comprising *A. bisporus GPD* promoter, 5'UTR sequences, ATG start codon and three naturally occurring introns.	14
SPR leader sequence (nucleotide sequence)	15
SPR leader sequence (amino acid sequence)	16
abst 1 terminator sequence	35
rafe terminator sequence	36
A. nidulans trpC terminator 3'UTR as engineered PCR product (800 nt) with introduced restriction sites *Bgl*II and *Kpn*I	37

19

Table 2: ASL-Argininosuccinate lyase (Genbank Accession no. AJ271691)

Intron	SEQ ID NO.
ASL intron 3	17
ASL intron 4	18
ASL intron 5	19
ASL intron 6	20

Table 3: CYPI-Cytochrome P450 I (Genbank Accession no. AJ271707)

Intron	SEQ ID NO.
CYPII intron 6	21
CYPII intron 7	22
CYPII intron 8	23
CYPII intron 9	24
CYPII intron 10	25

Table 4: SOD-Superoxide Dismutase (Genbank Accession no. AJ271694)

Intron	SEQ ID NO.
SOD intron 2	26
SOD intron	27

20

Table 5: LAP-Leucine Aminopeptidase (Genbank Accession no. AJ271690)

Intron	SEQ ID NO.
LAP intron 3	28
LAP intron 4	29
LAP intron 5	30
LAP intron 6	31
LAP intron 7	32

Table 6: MET-Metallothionein (Genbank Accession no. AJ271695)

Intron	SEQ ID NO.
MET intron 2	33
MET intron 3	34

21

CLAIMS:

1.   A filamentous fungus transformed with a heterologous sequence of DNA, the
fungus being capable of expressing the heterologous DNA, characterised in that the
heterologous DNA is under the control of a filamentous fungus transcription promoter
active substantially only during stage 1, or later, of the development of the fruiting body
of the fungus.

2.   A fungus according to claim 1, which is *A. bisporus*.

3.   A fungus according to claim 1 or 2, wherein the promoter is that of *abst1*.

4.   A fungus according to claim 1 or 2, wherein the promoter is that of *rafe*.

5.   A fungus according to claim 1 or 2, wherein the promoter is that of *mag2*.

6.   A fungus according to claim 1 or 2, wherein the promoter comprises the
sequence of SEQ ID NO 12, or a mutation or variant thereof, or a sequence which
hybridises thereto under conditions of at least 60% stringency.

7.   A fungus according to claim 1 or 2, wherein the promoter comprises the
sequence of SEQ ID NO 13, or a mutation or variant thereof, or a sequence which
hybridises thereto under conditions of at least 60% stringency.

8.   A fungus according to any preceding claim, wherein the DNA is operably linked
with a terminator comprising the sequence of SEQ ID NO 35, or a mutation or variant
thereof, or a sequence which hybridises thereto under conditions of at least 60%
stringency.

9.   A fungus according to any preceding claim, wherein the DNA is operably linked
with a terminator comprising the sequence of SEQ ID NO 36, or a mutation or variant
thereof, or a sequence which hybridises thereto under conditions of at least 60%
stringency.

22

10.   A fungus according to claim 1, wherein the DNA is operably linked with a promoter comprising the sequence of SEQ ID NO. 12 and a terminator comprising the sequence of SEQ ID NO 35, or a mutation or variant of either, or a sequence which hybridises thereto under conditions of at least 60% stringency.

11.   A fungus according to claim 1, wherein the DNA is operably linked with a promoter comprising the sequence of SEQ ID NO. 13 and a terminator comprising the sequence of SEQ ID NO 36, or a mutation or variant of either, or a sequence which hybridises thereto under conditions of at least 60% stringency.

12.   A fungus according to any preceding claim, wherein a selectable marker is linked with the heterologous DNA.

13.   A fungus according to any preceding claim, wherein the heterologous DNA is native DNA.

14.   A fungus according to any preceding claim, wherein the heterologous DNA is selected such as to affect characteristics of mushroom crop production.

15.   A fungus according to any preceding claim, wherein the heterologous DNA encodes: antibodies, including other diagnostic material; secondary metabolites, such as lectins, pesticidal compounds such as *Bacillus thuringiensis* toxin (Bt toxin); therapeutic compounds such as vaccines, steroids, heterocyclic organic compounds; biological macromolecules, such as interferon, endostatin and insulin; and medical enzymes, such as thrombolytics and cerebrosidases.

16.   A method for the production of a substance expressible by a DNA sequence, wherein the sequence is operably associated with a filamentous fungus transcription promoter active substantially only during stage 1, or later, of the development of the fruiting body of the fungus, the sequence and promoter being expressibly incorporated in a filamentous fungus, the fungus being cultured to fruition and the product being harvested.

23

17.   A method for the production of a substance expressible by a DNA sequence, wherein the sequence is operably associated with a filamentous fungus transcription promoter active substantially only during stage 1, or later, of the development of the fruiting body of the fungus, the sequence and promoter being expressibly incorporated in a filamentous fungus, the fungus being cultured to fruition and the product being harvested, and wherein the DNA and/or fungus is as defined in any of claims 1 to 15.
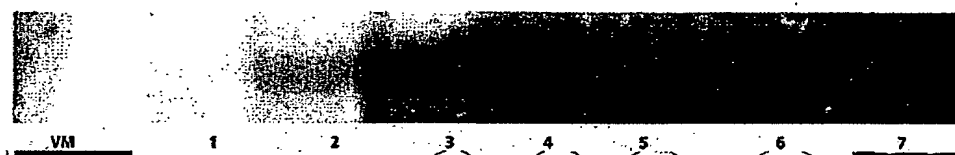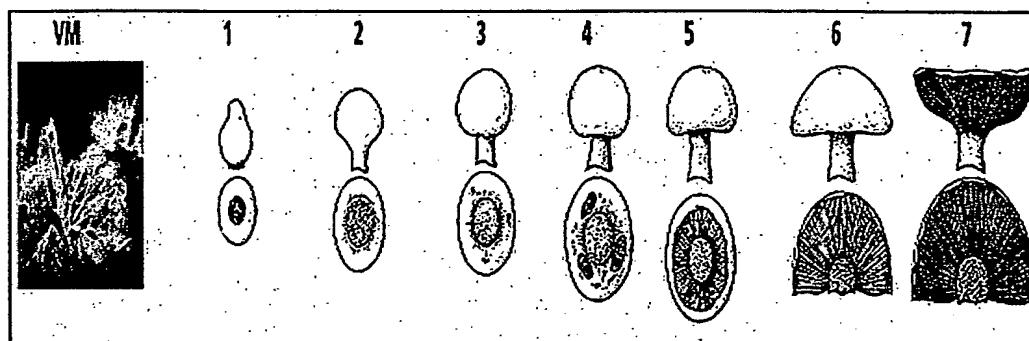
Fig.2A
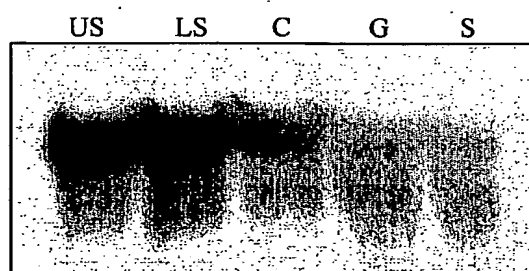
Fig.2B
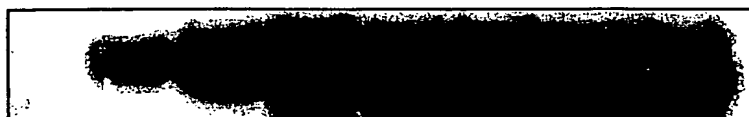
Fig.3A
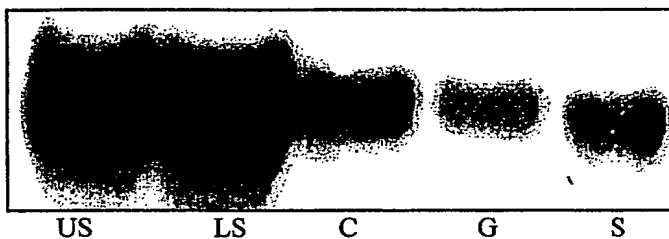
Fig.3B
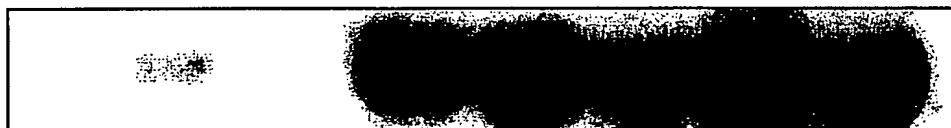
Fig.4A
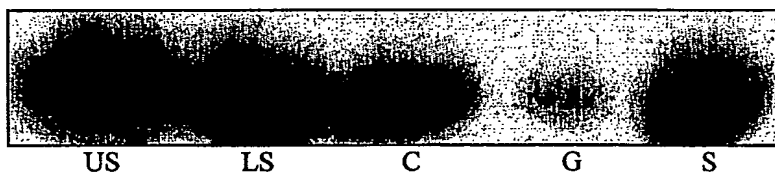
Fig.4B
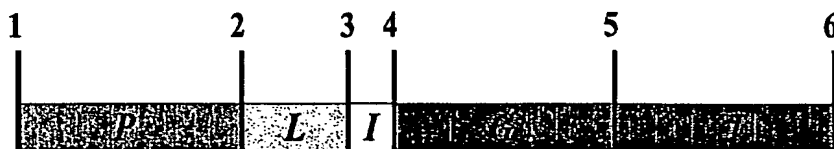
Fig.5
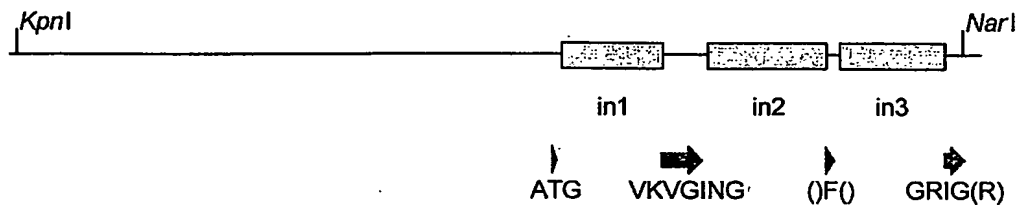
Fig.6

atgcatttct ctttgtcttt tgccaccctt gctctcttag tcgcttcggc tgttggt
M H F   S L S   F A T L   A L L   V A S   A V .G

Fig.7
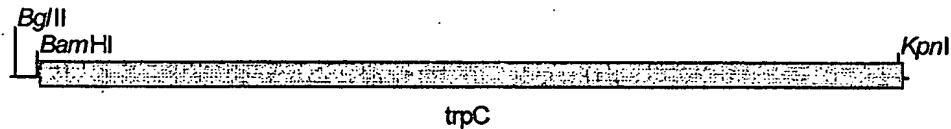
Fig.8
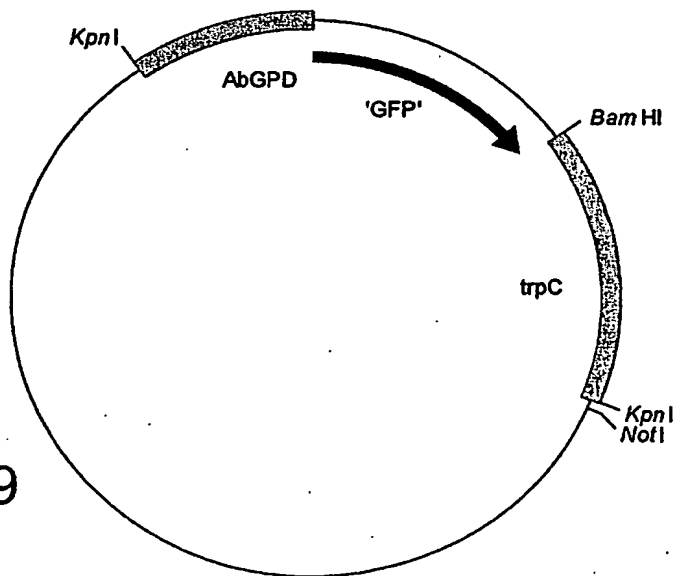
Fig.9

SEQUENCE LISTING

<110> Horticulture Research International

<120> SELECTIVE EXPRESSION IN FILAMENTOUS FUNGI

<130> WPP86709

<150> UK 0225390.4

<151> 2002-10-31

<160> 37

<170> PatentIn version 3.1

<210> 1

<211> 1736

<212> DNA

<213> Agaricus bisporus

<220>

<221> CDS

<222> (19)..(1674)

<223>

<400> 1
ggctgagctc tattcatc atg gcg tcg gaa cga cag att gaa gaa ctt ccc   51
                    Met Ala Ser Glu Arg Gln Ile Glu Glu Leu Pro
                    1   5   10

agt tat aag tac gcc tat att ttg acg gca tcg gct tgc ttg gga agt   99
Ser Tyr Lys Tyr Ala Tyr Ile Leu Thr Ala Ser Ala Cys Leu Gly Ser
            15   20   25

gtg ttt cac ggg tgg gat gta ggc ctt ata gga ggc ata ctc tcg cta   147
Val Phe His Gly Trp Asp Val Gly Leu Ile Gly Gly Ile Leu Ser Leu
        30   35   40

cgg tca ttt caa gaa tat ctc ggg atc aat aca aaa aat gcc gtc aag   195
Arg Ser Phe Gln Glu Tyr Leu Gly Ile Asn Thr Lys Asn Ala Val Lys
    45   50   55

aaa gcg att cta gac gga aac atc atc tct gtg ctc caa gcc gga tgt   243
Lys Ala Ile Leu Asp Gly Asn Ile Ile Ser Val Leu Gln Ala Gly Cys

```
Lys Ala Ile Leu Asp Gly Asn Ile Ile Ser Val Leu Gln Ala Gly Cys
60                  65                  70                  75

ttt ttt ggc gcg ctt gga acc gga tat ctc tct agt cga ttc ggc cga    291
Phe Phe Gly Ala Leu Gly Thr Gly Tyr Leu Ser Ser Arg Phe Gly Arg
                80                  85                  90

aga ccc tgt ctt att gca tct ggt att gtg tat ata act ggc ggt ttg    339
Arg Pro Cys Leu Ile Ala Ser Gly Ile Val Tyr Ile Thr Gly Gly Leu
            95                  100                 105

ctg caa tgc act gtc ggt ttg gga ccc tcg caa gct gct gct cta cac    387
Leu Gln Cys Thr Val Gly Leu Gly Pro Ser Gln Ala Ala Ala Leu His
        110                 115                 120

gtg ttc tat att ggc agg ttc att tct ggt atc ggt gtt ggg atg gtg    435
Val Phe Tyr Ile Gly Arg Phe Ile Ser Gly Ile Gly Val Gly Met Val
        125                 130                 135

tcc act ctc gtg cct ttg tat att tcg gag tgt gtc cct agg act ata    483
Ser Thr Leu Val Pro Leu Tyr Ile Ser Glu Cys Val Pro Arg Thr Ile
140                 145                 150                 155

cgc ggg cgc tgt act gga aca ctc caa ttt gcg act aac agt ggt ctg    531
Arg Gly Arg Cys Thr Gly Thr Leu Gln Phe Ala Thr Asn Ser Gly Leu
                160                 165                 170

atg ctg ggc ttc tgg gtc aac tac agc gtg tcg aaa aac gtg ccc ttt    579
Met Leu Gly Phe Trp Val Asn Tyr Ser Val Ser Lys Asn Val Pro Phe
                175                 180                 185

ggt gaa atg caa tgg cga att ccg tta att atc caa atg att ccg agc    627
Gly Glu Met Gln Trp Arg Ile Pro Leu Ile Ile Gln Met Ile Pro Ser
            190                 195                 200

ctc ttg ttc atc ata gcc atg ttt ttc caa cca gaa tcg ccg aga tgg    675
Leu Leu Phe Ile Ile Ala Met Phe Phe Gln Pro Glu Ser Pro Arg Trp
        205                 210                 215

ctt gtt gaa cac ggg aaa cac aag gaa gct gcg acg gta ctg gcg cgt    723
Leu Val Glu His Gly Lys His Lys Glu Ala Ala Thr Val Leu Ala Arg
220                 225                 230                 235

act ggc ggc aag gat gtt gat cat cct agt gtt gta cag aca ctg gag    771
Thr Gly Gly Lys Asp Val Asp His Pro Ser Val Val Gln Thr Leu Glu
            240                 245                 250

gag atc aag caa gaa ttt gtg gcg agt aaa caa cca tcg ttt tta aag    819
Glu Ile Lys Gln Glu Phe Val Ala Ser Lys Gln Pro Ser Phe Leu Lys
            255                 260                 265

cag att cgc ctg gtc ggt gaa tcg agg gct gtt gcc ctg agg tgc ttt    867
Gln Ile Arg Leu Val Gly Glu Ser Arg Ala Val Ala Leu Arg Cys Phe
            270                 275                 280

ata cca ccg cta gtg atg ttc ttc cag cag tgg acg ggt aca aat gcc    915
Ile Pro Pro Leu Val Met Phe Phe Gln Gln Trp Thr Gly Thr Asn Ala
            285                 290                 295

atc aac ctt tat agt ccc gaa gta ttc cgt cat ctt gga atc cat ggc    963
Ile Asn Leu Tyr Ser Pro Glu Val Phe Arg His Leu Gly Ile His Gly
300                 305                 310                 315

acc agc ggg gct ctc ttc gct act ggt gtt tat ggc gtg gtg aag gtt    1011
Thr Ser Gly Ala Leu Phe Ala Thr Gly Val Tyr Gly Val Val Lys Val
            320                 325                 330

gtt tca gtt gca ctt gcc ctc act ttt gct gtc gaa cgc ttt gga cgc    1059
```

```
Val Ser Val Ala Leu Ala Leu Thr Phe Ala Val Glu Arg Phe Gly Arg
        335             340             345

aag aga ggg ttg att ttt ggt ggt atc ggc caa gca ctt atg atg ttt      1107
Lys Arg Gly Leu Ile Phe Gly Gly Ile Gly Gln Ala Leu Met Met Phe
        350             355             360

tgg ttg gga ggt tat agt gcc acc cac caa gac ggt act gtc agt cct      1155
Trp Leu Gly Gly Tyr Ser Ala Thr His Gln Asp Gly Thr Val Ser Pro
        365             370             375

gcg agt cat gtt tcc att gtt gca ctc tac ttg tat ggt gca ttc ttc      1203
Ala Ser His Val Ser Ile Val Ala Leu Tyr Leu Tyr Gly Ala Phe Phe
380             385             390             395

tcc atg gga tgg ggc cca tta cca tgg gtc gtc gct gga gag gtt gca      1251
Ser Met Gly Trp Gly Pro Leu Pro Trp Val Val Ala Gly Glu Val Ala
            400             405             410

cct aac cat gtc cgc tcc ttc gcc ctc tcc atc gcc gtt gga act cat      1299
Pro Asn His Val Arg Ser Phe Ala Leu Ser Ile Ala Val Gly Thr His
        415             420             425

tgg ctc ttc ggg ttt gtg ata tca aaa gtg acg cca att atg ttg gac      1347
Trp Leu Phe Gly Phe Val Ile Ser Lys Val Thr Pro Ile Met Leu Asp
        430             435             440

cgt atc aaa tat ggc aca ttc cta ctc ttc gga ttc tgt tgc atg ata      1395
Arg Ile Lys Tyr Gly Thr Phe Leu Leu Phe Gly Phe Cys Cys Met Ile
        445             450             455

gta gcg aca tgg gct tat ttc tgt cta cct gag aca agt ggg ttc gct      1443
Val Ala Thr Trp Ala Tyr Phe Cys Leu Pro Glu Thr Ser Gly Phe Ala
460             465             470             475

ctg gag gac atc aaa tat ctg ttc gag cga gac gtc atc att cgt tca      1491
Leu Glu Asp Ile Lys Tyr Leu Phe Glu Arg Asp Val Ile Ile Arg Ser
            480             485             490

ttg cag gac gct ccc ggt gga aaa ata ttc ttg ggg ggg agg cgt gtg      1539
Leu Gln Asp Ala Pro Gly Gly Lys Ile Phe Leu Gly Gly Arg Arg Val
            495             500             505

gaa tct gta gct tcg ttg aaa gag agg cgc gtt gga gtc gct ggt gag      1587
Glu Ser Val Ala Ser Leu Lys Glu Arg Arg Val Gly Val Ala Gly Glu
            510             515             520

cag ggt gag aag ata act ggt cta aat tcg gaa ttg gaa gat gtt tcc      1635
Gln Gly Glu Lys Ile Thr Gly Leu Asn Ser Glu Leu Glu Asp Val Ser
        525             530             535

tca aaa aaa tca aca ttg aag gaa act tca tcc gtt tga tatatagtct      1684
Ser Lys Lys Ser Thr Leu Lys Glu Thr Ser Ser Val
540             545             550

ccaaattcta ttgtaatgcc attttcccaa ttcaaaaaaa aaaaaaaaaa aa          1736
```

<210> 2

<211> 551

<212> PRT

<213> Agaricus bisporus

<400> 2

Met Ala Ser Glu Arg Gln Ile Glu Glu Leu Pro Ser Tyr Lys Tyr Ala
1   5   10   15

Tyr Ile Leu Thr Ala Ser Ala Cys Leu Gly Ser Val Phe His Gly Trp
            20   25   30

Asp Val Gly Leu Ile Gly Gly Ile Leu Ser Leu Arg Ser Phe Gln Glu
        35   40   45

Tyr Leu Gly Ile Asn Thr Lys Asn Ala Val Lys Lys Ala Ile Leu Asp
    50   55   60

Gly Asn Ile Ile Ser Val Leu Gln Ala Gly Cys Phe Phe Gly Ala Leu
65   70   75   80

Gly Thr Gly Tyr Leu Ser Ser Arg Phe Gly Arg Arg Pro Cys Leu Ile
            85   90   95

Ala Ser Gly Ile Val Tyr Ile Thr Gly Gly Leu Leu Gln Cys Thr Val
            100   105   110

Gly Leu Gly Pro Ser Gln Ala Ala Ala Leu His Val Phe Tyr Ile Gly
        115   120   125

Arg Phe Ile Ser Gly Ile Gly Val Gly Met Val Ser Thr Leu Val Pro
    130   135   140

Leu Tyr Ile Ser Glu Cys Val Pro Arg Thr Ile Arg Gly Arg Cys Thr
145   150   155   160

Gly Thr Leu Gln Phe Ala Thr Asn Ser Gly Leu Met Leu Gly Phe Trp
            165   170   175

Val Asn Tyr Ser Val Ser Lys Asn Val Pro Phe Gly Glu Met Gln Trp
            180   185   190

Arg Ile Pro Leu Ile Ile Gln Met Ile Pro Ser Leu Leu Phe Ile Ile
        195   200   205

Ala Met Phe Phe Gln Pro Glu Ser Pro Arg Trp Leu Val Glu His Gly
    210   215   220

Lys His Lys Glu Ala Ala Thr Val Leu Ala Arg Thr Gly Gly Lys Asp
225   230   235   240

Val Asp His Pro Ser Val Val Gln Thr Leu Glu Glu Ile Lys Gln Glu
            245   250   255

Phe Val Ala Ser Lys Gln Pro Ser Phe Leu Lys Gln Ile Arg Leu Val
            260   265   270

```
Gly Glu Ser Arg Ala Val Ala Leu Arg Cys Phe Ile Pro Pro Leu Val
        275             280             285

Met Phe Phe Gln Gln Trp Thr Gly Thr Asn Ala Ile Asn Leu Tyr Ser
        290             295             300

Pro Glu Val Phe Arg His Leu Gly Ile His Gly Thr Ser Gly Ala Leu
305             310             315             320

Phe Ala Thr Gly Val Tyr Gly Val Val Lys Val Val Ser Val Ala Leu
            325             330             335

Ala Leu Thr Phe Ala Val Glu Arg Phe Gly Arg Lys Arg Gly Leu Ile
        340             345             350

Phe Gly Gly Ile Gly Gln Ala Leu Met Met Phe Trp Leu Gly Gly Tyr
        355             360             365

Ser Ala Thr His Gln Asp Gly Thr Val Ser Pro Ala Ser His Val Ser
        370             375             380

Ile Val Ala Leu Tyr Leu Tyr Gly Ala Phe Phe Ser Met Gly Trp Gly
385             390             395             400

Pro Leu Pro Trp Val Val Ala Gly Glu Val Ala Pro Asn His Val Arg
            405             410             415

Ser Phe Ala Leu Ser Ile Ala Val Gly Thr His Trp Leu Phe Gly Phe
        420             425             430

Val Ile Ser Lys Val Thr Pro Ile Met Leu Asp Arg Ile Lys Tyr Gly
        435             440             445

Thr Phe Leu Leu Phe Gly Phe Cys Cys Met Ile Val Ala Thr Trp Ala
        450             455             460

Tyr Phe Cys Leu Pro Glu Thr Ser Gly Phe Ala Leu Glu Asp Ile Lys
465             470             475             480

Tyr Leu Phe Glu Arg Asp Val Ile Ile Arg Ser Leu Gln Asp Ala Pro
            485             490             495

Gly Gly Lys Ile Phe Leu Gly Gly Arg Arg Val Glu Ser Val Ala Ser
            500             505             510

Leu Lys Glu Arg Arg Val Gly Val Ala Gly Glu Gln Gly Glu Lys Ile
        515             520             525

Thr Gly Leu Asn Ser Glu Leu Glu Asp Val Ser Ser Lys Lys Ser Thr
        530             535             540
```

```
Leu Lys Glu Thr Ser Ser Val
545                 550


<210>  3

<211>  3638

<212>  DNA

<213>  Agaricus bisporus


<220>

<221>  misc_feature

<222>  (3)..(3)

<223>  n is an unknown nucleotide


<220>

<221>  misc_feature

<222>  (14)..(14)

<223>  n is an unknown nucleotide


<220>

<221>  misc_feature

<222>  (510)..(510)

<223>  n is an unknown nucleotide


<400>  3
gtncgatggg ttcntctggg ttaagttgca cgacgctttc cttttctttt tatggcctgt      60

ctgccctttt aacgctttat ctttcggcag ccatggatgt ccttcgtcac cgtattatca     120

ctcttaatcg tggtggtgag cacatggaaa ggttcattat cgtcatccca tgacgcggtg     180

caaaattcgt cattcagagt ggaaccgata ctaggagagg attttgaaaa ggctatcgta     240

tcttccgatc cattcactcc aacgacatct atatcgggtt ctcaagctac gataaccgtt     300

cctcacgcat cacattcttc tctgattgta atgaccgaag accgcggttc cacccctaat     360

attcgaagag cttgttgcca atgtcaaaaa aggcttttgg gcaaaccggg atggcttatc     420

gagcctccag ccccagcagc taacatcggg cagagaagga aaattcatcg ggcggttgaa     480

ttatcaccgt ttggttcctg agtcatctgn ragatgtacg cagatggtga taccgtgttt     540

gattggcgcc gttggagaag aactatatta ttcgatggat tttttgttcg agtttgacac     600

agagacagag atgatagagg tttgctattg atgtagcaaa ggatcatttg acgatggcgc     660
```

```
ataggggcgat ggttatcttt atgtctggaa ttataatatg tattgttccc cacttttctt    720

ttatatttat taatactaat tggaagtttc agttgttgga tgagcaaagt tggtgcagat    780

agaaactaga attcggattc ccatatctga ggtacctttt ccttccgctg gcaatcctgg    840

ccacttcgac gtggtgacgc agagggcgcg tgctattgtt agcacatgcc atawggatcg    900

acgttgcctc tcgtacttcg cgcctaggct cgctcatgcc tcgatgcatc ttttcaattc    960

gggcgttgcg tctcccaggt gcctgttaaa agggcgaact ttagtgtaat tgtactaaca   1020

cagtccctcg ggctgagctc tattcatcat ggcgtcggaa cgacagattg aagaacttcc   1080

cagttataag tacgcctata ttttgacggc atcggcttgc ttgggaagtg tgtttcacgg   1140

gtgggatgtg caagtatcct ttwgacgctg gctcatccct tgtttgctga tggttaatgt   1200

gtagaggcct tataggaggc atactctcgc tacggtcatt tcaagaatat ctcgggatca   1260

atacaaaaaa tgccgtcaag aaagcgattc tagacggaaa catcatctct gtgctccaag   1320

ccggatgttt tgtaagtcgt agcactcgtt cgaccagctt actttctttt actaacaacg   1380

tgttgaccta gttttggcgc gcttggaacc ggatatctct ctagtcgatt cggccgaaga   1440

ccctgtctta ttgcatctgg tattgtgtat ataactggcg gtttgctgca atgcactgtc   1500

ggtttgggac cctcgcaagc tgctgctcta cacgtgttct atattggcag gttcatttct   1560

ggtatcggtg ttgggatggt gtccactctc gtgcctttgt atatttcgga gtgtgtccct   1620

aggactatac gcgggcgctg tactggaaca ctccaatttg cgactaacag tggtctgatg   1680

ctgggctgta agtgtacctt ttgttatgct ccggggacga tactaaaagt agtctgtagt   1740

ctgggtcaac tacagcgtgt cgaaaaacgt gccctttggt gaaatgcaat ggcgaattcc   1800

gttaattatc cagtacgtta ttggatgcaa agtgaagcat atgctgaaat tagacttgtg   1860

tcgtactttc cagaatgatt ccgagcctct tgttcatcat agccatgttt ttccaaccag   1920

aatcgccgag atggcttgtt gaacacggga aacacaagga agctgcgacg gtactggcgc   1980

gtactggcgg caaggatgtt gatcatccta gtgttgtaca gacactggag gagatcaagc   2040

aagaatttgt ggcgagtaaa caaccatcgt ttttaaagca gattcgcctg gtcggtgaat   2100

cgagggctgt tgccytgagg wgctttatac caccgctagt gatgttcttc cagcagtgga   2160

cgggtacaaa tgccatcaac ctttatagtc ccgaagtatt ccgtcatctt ggaatccatg   2220

gcaycagckr kgmwstcttc sctamtsgtg tttatkgmgt ggtrmaggtt gtttcagttg   2280

cacttgccct cacttttgct gtcgaacgct ttggacgcaa gagagggttg atttttggtg   2340

gtatcggcca agcacttatg atgttttggt tgggaggtta tagtgccacc caccaagacg   2400

gtactgtcag tcctgcgagt catgtttcca ttgttgcact ctacttgtat ggtgcattct   2460

tctccactga agaatggttc ctgttabcat gggtcgtcgc tggagaggtt gcacctaacc   2520

atgtccgctc cttcgccctc tccatcgccg ttggaactca ttggctcttc gggtttgtga   2580

tatcaaaagt gacgccaatt atgttggacc gtatcaaata tggcacattc ctactcttcg   2640

gattctgttg catgatagta gcgacatggg cttatttctg tctacctgag acaagtgggt   2700
```

```
tcgctctgga ggacatcaaa tatctgttcg agcgagacgt catcattcgt tcattgcagg    2760

acgctcccgg tggaaaaata ttcttggggg ggaggcgtgt ggaatctgta gcttcgttga    2820

aagagaggcg cgttggagtc gctggtgagc agggtgagaa gataactggt ctaaattcgg    2880

aattggaaga tgtttcctca aaaaaatcaa cattgaagga aacttcatcc gtttgatata    2940

tagtctccaa attctattgt aatgccattt tcccaattca aaaggacccg ctctcgaacc    3000

gggtcagatg caattttggt cagcaatggt ttatgttgtt tccccgtaag tatgcactag    3060

aagagaacaa aacgtcacta tttgctcaat gcaggatgca cctggcgaga taatattctt    3120

gcggtgaagt cgaacaacgt ctgtagtcct gtaaaaatat acagtgagta gagggatgat    3180

gccgatgtgg aaggaagcaa ccgattacga ttcggatttg gccatgagac ggccgctctt    3240

gaacagacca acaatatccc tttaaattta atacagaatt actcaatatg cttccaagta    3300

tttcgattcc tcgaattccg tctggccgcg gcatgagcat ggacaggcgg acagaagagg    3360

ctatcgttgt attgcttcat caagcgaccc tgactagtga cttcaggcat gatcatgcgc    3420

ttagcaatct gtcccttcaa gtcgagtccc cgaattcaac agctttcaac aagtcgtgat    3480

tatttgaccc ccgactggaa tcaaattggc tcttcaaatt tcaaaccttc aatgcttcat    3540

gcttcatgcg tcatgacgca agctgtcaat tttcattttc cagttcggtc ccattctcac    3600

tctcgcccct cstaatgtct tccagaaaat ggcttccc                            3638
```

<210> 4

<211> 1033

<212> DNA

<213> Agaricus bisporus

<220>

<221> misc_feature

<222> (495)..(495)

<223> n is an unknown nucleotide

<400> 4
```
ctgggttaag ttgcacgacg ctttcctttt tcttttatgg cctgtctgcc cttttaacgc      60

tttatctttc ggcagccatg gatgtccttc gtcaccgtat tatcactctt aatcgtggtg     120

gtgagcacat ggaaaggttc attatcgtca tcccatgacg cggtgcaaaa ttcgtcattc     180

agagtggaac cgatactagg agaggatttt gaaaaggcta tcgtatcttc cgatccattc     240

actccaacga catctatatc gggttctcaa gctacgataa ccgttcctca cgcatcacat     300

tcttctctga ttgtaatgac cgaagaccgc ggttccaccc ctaatattcg aagagcttgt     360

tgccaatgtc aaaaaaggct tttgggcaaa ccgggatggc ttatcgagcc tccagcccca     420
```

```
gcagctaaca tcgggcagag aaggaaaatt catcgggcgg ttgaattatc accgtttggt    480

tcctgagtca tctgnragat gtacgcagat ggtgataccg tgtttgattg gcgccgttgg    540

agaagaacta tattattcga tggattttttt gttcgagttt gacacagaga cagagatgat    600

agaggtttgc tattgatgta gcaaaggatc atttgacgat ggcgcatagg gcgatggtta    660

tctttatgtc tggaattata atatgtattg ttccccactt ttcttttata tttattaata    720

ctaattggaa gtttcagttg ttggatgagc aaagttggtg cagatagaaa ctagaattcg    780

gattcccata tctgaggtac cttttccttc cgctggcaat cctggccact tcgacgtggt    840

gacgcagagg gcgcgtgcta ttgttagcac atgccatawg gatcgacgtt gcctctcgta    900

cttcgcgcct aggctcgctc atgcctcgat gcatctttttc aattcgggcg ttgcgtctcc    960

caggtgcctg ttaaaagggc gaactttagt gtaattgtac taacacagtc cctcgggctg   1020

agctctattc atc                                                      1033
```

<210> 5

<211> 702

<212> DNA

<213> Agaricus bisporus

<400> 5

```
tatatagtct ccaaattcta ttgtaatgcc attttcccaa ttcaaaagga cccgctctcg     60

aaccgggtca gatgcaattt tggtcagcaa tggtttatgt tgtttccccg taagtatgca    120

ctagaagaga acaaaacgtc actatttgct caatgcagga tgcacctggc gagataatat    180

tcttgcggtg aagtcgaaca acgtctgtag tcctgtaaaa atatacagtg agtagaggga    240

tgatgccgat gtggaaggaa gcaaccgatt acgattcgga tttggccatg agacggccgc    300

tcttgaacag accaacaata tccctttaaa tttaatacag aattactcaa tatgcttcca    360

agtatttcga ttcctcgaat tccgtctggc cgcggcatga gcatggacag gcggacagaa    420

gaggctatcg ttgtattgct tcatcaagcg accctgacta gtgacttcag gcatgatcat    480

gcgcttagca atctgtccct tcaagtcgag tccccgaatt caacagcttt caacaagtcg    540

tgattatttg acccccgact ggaatcaaat tggctcttca aatttcaaac cttcaatgct    600

tcatgcttca tgcgtcatga cgcaagctgt caattttcat tttccagttc ggtcccattc    660

tcactctcgc ccctcstaat gtcttccaga aaatggcttc cc                       702
```

<210> 6

<211> 689

<212> DNA

<213> Agaricus bisporus

<220>

<221> CDS

<222> (147)..(521)

<223>

<400> 6
cgcctacttt tataccaacc ccaaatccaa aggttgaaaa aaaaatttcg acaaggattt   60

atatatccat ccatccgcga cactttcccg tttgattcta tcccttagtc tttccttctc   120

cccctttcct tcttcacttc accttg ctc taa ccg aaa gta aac ctt tcc gcg   173
                             Leu   Pro Lys Val Asn Leu Ser Ala
                             1   5

atg aaa ttc tcc aat tct cta tcc gct ctc ctc gta tcc gca aat ctc   221
Met Lys Phe Ser Asn Ser Leu Ser Ala Leu Leu Val Ser Ala Asn Leu
    10   15   20

atg ttg gca gcg aag gcc tac aaa gga gat gcc acc ttt tat gat cct   269
Met Leu Ala Ala Lys Ala Tyr Lys Gly Asp Ala Thr Phe Tyr Asp Pro
25   30   35   40

ggt ctg gga gct tgt ggc cat acg aat cag gct cat gaa ctt gtc gtt   317
Gly Leu Gly Ala Cys Gly His Thr Asn Gln Ala His Glu Leu Val Val
            45   50   55

gcc ctt cca tca gcc aaa tac ggc agc gga gac cat tgt tcc aag cat   365
Ala Leu Pro Ser Ala Lys Tyr Gly Ser Gly Asp His Cys Ser Lys His
            60   65   70

gtc ggc atc cac tac aaa ggc aaa tac gtg aaa gcc aaa gta gtc gac   413
Val Gly Ile His Tyr Lys Gly Lys Tyr Val Lys Ala Lys Val Val Asp
            75   80   85

aaa tgt ccc ggt tgt ggt tcg aac gat tta gac atc tca cca acc gca   461
Lys Cys Pro Gly Cys Gly Ser Asn Asp Leu Asp Ile Ser Pro Thr Ala
    90   95   100

ttc tct cac tta gcc agt caa gac ctc ggc cgt atc aaa gta gat tgg   509
Phe Ser His Leu Ala Ser Gln Asp Leu Gly Arg Ile Lys Val Asp Trp
105   110   115   120

gaa ttt ctc tga tatcccattt tcaatccctt acacgaaatc tgtatttgta   561
Glu Phe Leu

gaagaaagtc atgacgttat atagatcact tacatagatc ttcaggtttt cgtagatcga   621

cgaccgacgc tcttaaatat ttatttcccg tttttcygtt tttgtttaa aaaaaaaaaa   681

aaaaaaaa   689

<210> 7

<211> 122

<212> PRT

<213> Agaricus bisporus

10/24

<400> 7

Pro Lys Val Asn Leu Ser Ala Met Lys Phe Ser Asn Ser Leu Ser Ala
1   5   10   15

Leu Leu Val Ser Ala Asn Leu Met Leu Ala Ala Lys Ala Tyr Lys Gly
            20   25   30

Asp Ala Thr Phe Tyr Asp Pro Gly Leu Gly Ala Cys Gly His Thr Asn
        35   40   45

Gln Ala His Glu Leu Val Val Ala Leu Pro Ser Ala Lys Tyr Gly Ser
    50   55   60

Gly Asp His Cys Ser Lys His Val Gly Ile His Tyr Lys Gly Lys Tyr
65   70   75   80

Val Lys Ala Lys Val Val Asp Lys Cys Pro Gly Cys Gly Ser Asn Asp
                85   90   95

Leu Asp Ile Ser Pro Thr Ala Phe Ser His Leu Ala Ser Gln Asp Leu
            100   105   110

Gly Arg Ile Lys Val Asp Trp Glu Phe Leu
            115   120

<210> 8

<211> 2782

<212> DNA

<213> Agaricus bisporus

<220>

<221> misc_feature

<222> (2774)..(2774)

<223> n is an unknown nucleotide

<400> 8
ctgcgaattg gcataagcac ttraactttc gtcttcctca ctctcttcag gagattgaga   60

ctgcatcggt gcaagcgagg gttgccgtac cgccctttga gaccccgaaa caagggattc   120

gtcaagagca tcaagaagac tagaacgacc ggctgttttt ccacccgaca tcatagcaca   180

aactgtcata aacccgtgtt caaagggaa aaacaggcag agagaaggaa gggacgcgtc   240

gcgatgaaat gctcattaac ctgaatgaca aacttccgcw aataacaatt taattaaaaa   300

taaaaatcac gaggtgacaa acaggggtgt ttacctccat tcgactgcat cctggctctt   360

```
cccttccata gaactgtcgt ccatgcatgc accgctagca tcgcactgct ctgactcgca    420

taaccttaaa acgcgtggac cccctgttcg gacggccggt tcaggatccg gggctcagga    480

cacagtaaaa tcacaaaaac tcatactttg agagatatga cttctcgact tgcgccttcg    540

atggacggac aaattatccc caggtaccgg atctgtgaca ccgaattagt gcgcgatatt    600

atatatgact tttgacgggc gtctcatacg accgctcaag tccttgggga tggagaatgt    660

cacctcctgg tccaccgggc ccagagcatt acccggtcat taatctagcg cttcttgcat    720

gcactcctgc atgatcaccc cacgcggccg cgttttatcg gacatataag gaacaagatt    780

ccataggtag tggatcccct actccacctc ccgcctactt ttataccaac cccaaatcca    840

aaggttgaaa aaaaatttc gacaaggatt tatatatcca tccatccgcg acactttctc     900

gtttgattct atcccttagt ctttccttct ccccctttcc ttcttcactt caccttgctc    960

taaccgaaag taaacctttc cgcgatgaaa ttctccaatt ctctatccgc tctcctcgta   1020

tccgcgaatc tcatgttggc agcgaaggcc tacaaaggag atggtcagtt atgacagtgt   1080

cgcccatcct atcgtcaatg atgctgattt tgttctccca ccttttagc caccttttat    1140

gatcctggtg tacgtttttt ttacactttt caccaagatt cctcacccac tggtttcaaa   1200

attagctggg agcttgtggc catacgaatc aggctcatga acttgttgtt gcccttccat   1260

cagccaaata cggcaacgga gaccattgtt ccaagcatgt cggcatccac tgtacgtttc   1320

ctccttctcc ccacttctca aaaatcaaaa ttttactcat tgtaacaaga caaaggcaaa   1380

tacgtgaaag tcaaagtagt cgacaaatgt cccggctgtg gttcgaacga tctagacatc   1440

tcaccaaccg cattctctca gttagccagc caagacctcg gccgtatcaa agtagattgg   1500

gaatttatct aatatcccat ttcaatccct tacacgaaat ctgtacttgt agtcttagaa   1560

gaaagtcatg acgttatata gatcacttac atagatcttc aggttttcgt agatcgacga   1620

ccgacgctct taaatattta tttacagttt ttctgttttt tgttttattg tcgcttggat   1680

ataaggtggt atactttgat atgattgcct acacacatat atcaacacag ttttagttat   1740

atcaacatca aaacatcagt caaggaaaac aaagagcgaa cgataaacat cagcacaagt   1800

atgtcagatt atggtccaag aacgcgaaaa gaagttcgca aaagaacaga acactatcga   1860

aaagtgcaga tacataggtc acacaattaa cgacttcccg gaatagttcc ctccaacctc   1920

ttatcgcgac tactagcacc aacggtaaca ccaaaagtac cttcaggcct cctccatccc   1980

tgtgcattca catcccaaat actcaaatca tacctcgaca aggtcatttt tacattccta   2040

gtctctccag ggccaatcgg tacagagtcg aaaccgcgta gcacggaagg aggttctcca   2100

gcagattcag ggaagttaat gtagagttgg ggagactcgg cacccaaaag tcgaccggta   2160

ttcttgacgt tgaaggaaac ctcgtacaaa ggacgatgga gcctggaaac cggtcagcgt   2220

gaaattgttg rttagagaac cctcggcgct taccaagaag cgatagaacc accttcaact   2280

tgcgggcttg catgtccttg atcccattcc ctagccgcat caacttctct acctttgacg   2340

tgcactttct taatagatat cttcgagtaa tcaaacttcg tgtaactcaa gccaaaacca   2400
```

12/24

```
aattcaaacc gaggctcaat gccattctgc aagagaaaat tcagtaagcg aacgtttccg    2460

agagaatcaa gcagcactgg cttacagcgt cgaagtgccg gtaatcgatt tgcaaacgat    2520

catcgtatgg tatagaaaga atatctcgaa taccatctcc ggggataact tgggccgaat    2580

aatcttcaat ccgtttggca atggtatatg gaagccttcc ggatgggttc caatctccat    2640

ataggacatc ggtcaacgaa tttcccgctt ctgtcccggt gccccagccc aaagaacctg    2700

gaaatagtca acaccggcgt tcatcacaca agtagaagat acgaacagca gtgacattag    2760

gatgattgat ccanggttca ag                                            2782
```

<210> 9

<211> 984

<212> DNA

<213> Agaricus bisporus

<400> 9
```
ctgcgaattg gcataagcac ttraactttc gtcttcctca ctctcttcag gagattgaga     60

ctgcatcggt gcaagcgagg gttgccgtac cgccctttga gaccccgaaa caagggattc    120

gtcaagagca tcaagaagac tagaacgacc ggctgttttt ccacccgaca tcatagcaca    180

aactgtcata aacccgtgtt caaaggggaa aaacaggcag agagaaggaa gggacgcgtc    240

gcgatgaaat gctcattaac ctgaatgaca aacttccgcw aataacaatt taattaaaaa    300

taaaaatcac gaggtgacaa acaggggtgt ttacctccat tcgactgcat cctggctctt    360

cccttccata gaactgtcgt ccatgcatgc accgctagca tcgcactgct ctgactcgca    420

taaccttaaa acgcgtggac cccctgttcg gacggccggt tcaggatccg gggctcagga    480

cacagtaaaa tcacaaaaac tcatactttg agagatatga cttctcgact tgcgccttcg    540

atggacggac aaattatccc caggtaccgg atctgtgaca ccgaattagt gcgcgatatt    600

atatatgact tttgacgggc gtctcatacg accgctcaag tccttgggga tggagaatgt    660

cacctcctgg tccaccgggc ccagagcatt acccggtcat taatctagcg cttcttgcat    720

gcactcctgc atgatcaccc cacgcggccg cgtttatcg gacatataag gaacaagatt    780

ccataggtag tggatccctt actccacctc ccgcctactt ttataccaac cccaaatcca    840

aaggttgaaa aaaaatttc gacaaggatt tatatatcca tccatccgcg acactttctc    900

gtttgattct atcccttagt ctttccttct cccccttttcc ttcttcactt caccttgctc    960

taaccgaaag taaaccttc cgcg                                            984
```

<210> 10

<211> 1270

<212> DNA

<213> Agaricus bisporus

<220>

<221> misc_feature

<222> (1262)..(1262)

<223> n is an unknown nucleotide

<400> 10

```
tatcccattt caatccctta cacgaaatct gtacttgtag tcttagaaga aagtcatgac      60

gttatataga tcacttacat agatcttcag gttttcgtag atcgacgacc gacgctctta     120

aatatttatt tacagttttt ctgttttttg ttttattgtc gcttggatat aaggtggtat     180

actttgatat gattgcctac acacatatat caacacagtt ttagttatat caacatcaaa     240

acatcagtca aggaaaacaa agagcgaacg ataaacatca gcacaagtat gtcagattat     300

ggtccaagaa cgcgaaaaga agttcgcaaa agaacagaac actatcgaaa agtgcagata     360

cataggtcac acaattaacg acttcccgga atagttccct ccaacctctt atcgcgacta     420

ctagcaccaa cggtaacacc aaaagtacct tcaggcctcc tccatccctg tgcattcaca     480

tcccaaatac tcaaatcata cctcgacaag gtcattttta cattcctagt ctctccaggg     540

ccaatcggta cagagtcgaa accgcgtagc acggaaggag gttctccagc agattcaggg     600

aagttaatgt agagttgggg agactcggca cccaaaagtc gaccggtatt cttgacgttg     660

aaggaaacct cgtacaaagg acgatggagc ctggaaaccg gtcagcgtga aattgttgrt     720

tagagaaccc tcggcgctta ccaagaagcg atagaaccac cttcaacttg cgggcttgca     780

tgtccttgat cccattccct agccgcatca acttctctac ctttgacgtg cactttctta     840

atagatatct tcgagtaatc aaacttcgtg taactcaagc caaaaccaaa ttcaaaccga     900

ggctcaatgc cattctgcaa gagaaaattc agtaagcgaa cgtttccgag agaatcaagc     960

agcactggct tacagcgtcg aagtgccggt aatcgatttg caaacgatca tcgtatggta    1020

tagaaagaat atctcgaata ccatctccgg ggataacttg ggccgaataa tcttcaatcc    1080

gtttggcaat ggtatatgga agccttccgg atgggttcca atctccatat aggacatcgg    1140

tcaacgaatt tcccgcttct gtcccggtgc cccagcccaa agaacctgga aatagtcaac    1200

accggcgttc atcacacaag tagaagatac gaacagcagt gacattagga tgattgatcc    1260

anggttcaag                                                          1270
```

<210> 11

<211> 835

<212> DNA

<213> Agaricus bisporus

```
<400>  11
ccttgccgtt ttccagaagc tgcgaccagt cctcggagga gggggaacta aagatgtcga      60

aagcagcagg gtgtgtgaca tagcgggaag aaagggtgat gatgatgaga acgaggggta     120

cgatgaggac ggcccatctg aattgacggc caacgcgacg tttccggtcg tgagaggaca     180

tggcaaagga gacgggggga ggggcgaggg tggcggagga ggtgctcgtg ccgaattcgg     240

cacgagctca ccatgaaatt cgcaactgct ctcctcgcct gccttactgc tgctgctagc     300

gctcarcgcg tcctcatcgg atcccctcct gaccaagcaa atctttctgc tggccagaac     360

actacgattc aaattgtact cccgaatttc caatcgtctt cgcaagaagt tgcggtagtg     420

cttgggatca cgtcctgcgc cgctgctccc tgccctgctc cagccgatac gatgggtcgt     480

atcctttaca gcggtcattt caacccgcag agagatcctg caatgcccgc aatgcaagcc     540

tacgaaaatt tcacggtctt cttgccggag aacctgccta agggcgcggc gcagattaac     600

gtttaccatg tcgcacttat cggggccggt ctcatgccat ggaacgagac gttgtccacc     660

acagctttga ttcagtaatt ccatcaggat ttgaaatgga cctttagtag tttactgttt     720

tgctatcgaa cgattcgrat aattacctga gatcaggtcg gtgactgagg cccgtcggag     780

tgctaccata atggcataat aaaattatac tcagctgaaa aaaaaaaaa aaaaa          835


<210>  12

<211>  770

<212>  DNA

<213>  Agaricus bisporus


<400>  12
tcgggttctc aagctacgat aaccgttcct cacgcatcac attcttctct gattgtaatg      60

accgaagacc gcggttccac ccctaatatt cgaagagctt gttgccaatg tcaaaaaagg     120

cttttgggca aaccgggatg gcttatcgag cctccagccc cagcagctaa catcgggcag     180

agaaggaaaa tcatcggcgt tgaattatca ccgtttggtt cctgagtcat ctggagatgt     240

acgcagatgg tgataccgtg tttgattggc gccgttggag aagaactata ttattcgatg     300

gatttttgt tcgagtttga cacagagaca gagatgatag aggtttgcta ttgatgtagc     360

aaaggatcat ttgacgatgg cgcatagggc gatggttatc tttatgtctg gaattataat     420

atgtattgtt ccccactttt cttttatatt tattaatact aattggaagt ttcagttgtt     480

ggatgagcaa agttggtgca gatagaaact agaattcgga ttcccatatc tgaggtacct     540

tttccttccg ctggcaatcc tggccacttc gacgtggtga cgcagagggc gcgtgctatt     600

gttagcacat gccatatgga tcgacgttgc ctctcgtact tcgcgcctag gctcgctcat     660

gcctcgatgc atctttcaat tcgggcgttg cgtctcccag gtgcctgtta aaagggcgaa     720

ctttagtgta attgtactaa cacagtccct cgggctgagc tctattcatc                 770
```

```
<210>  13
<211>  703
<212>  DNA
<213>  Agaricus bisporus


<400>  13
ataacaattt aattaaaaat aaaaatcacg aggtgacaaa cagggtgtt tacctccatt      60
cgactgcatc ctggctcttc ccttccatag aactgtcgtc catgcatgca ccgctagcat     120
cgcactgctc tgactcgcat aaccttaaaa cgcgtggacc ccctgttcgg acggccggtt     180
caggatccgg ggctcaggac acagtaaaat cacaaaaact catactttga gagatatgac     240
ttctcgactt gcgccttcga tggacggaca aattatcccc aggtaccgga tctgtgacac     300
cgaattagtg cgcgatatta tatatgactt ttgacgggcg tctcatacga ccgctcaagt     360
ccttggggat ggagaatgtc acctcctggt ccaccgggcc cagagcatta cccggtcatt     420
aatctagcgc ttcttgcatg cactcctgca tgatcacccc acgcggccgc gttttatcgg     480
acatataagg aacaagattc cataggtagt ggatccccta ctccacctcc cgcctacttt     540
tataccaacc ccaaatccaa aggttgaaaa aaaaatttcg acaaggattt atatatccat     600
ccatccgcga cactttctcg tttgattcta tcccttagtc tttccttctc cccctttcct     660
tcttcacttc accttgctct aaccgaaagt aaacctttcc gcg                       703
```

```
<210>  14
<211>  486
<212>  DNA
<213>  Agaricus bisporus


<220>
<221>  misc_feature
<222>  (4)..(9)
<223>  Restriction site for KpnI


<220>
<221>  misc_feature
<222>  (477)..(482)
<223>  Restriction site for NarI
```

<220>

<221> Intron

<222> (277)..(328)

<223>

<220>

<221> Intron

<222> (349)..(408)

<223>

<220>

<221> Intron

<222> (415)..(468)

<223>

<400> 14
```
cggggtaccg aggtccgcaa gtagattgaa agttcagtac gttttttaaca atagagcatt     60

ctcgaggctt gcgtcattct gtgtcaggct agcagtttat aagcgttgag gatctagagc    120

tgctgtttcc gcgtctcgaa tgttctcggt gtttaggggt tagcaatctg atatgataat    180

aatttgtgat gacatcgata gtacaaaaac cccaattccg gtcacatcca ccatctccgt    240

tttctcccat ctacacacaa caagcttatc gccatggttt gtctctcgct tgcataccat    300

ccagcagctc actgatgtcg acttgtaggt taaagttgga atcaacgggt aagtgttttt    360

gtcgtcgcgc tgtggttccg gatcatctca gactttgggt gtcttgcagt ttcggtgagt    420

gaccaccctg cattctggct atatgcgtga tactgaccat cgctcaaggt cgtatcggcg    480

ccggcc                                                              486
```

<210> 15

<211> 57

<212> DNA

<213> Agaricus bisporus

<220>

<221> CDS

<222> (1)..(57)

<223>

```
<400> 15
atg cat ttc tct ttg tct ttt gcc acc ctt gct ctc tta gtc gct tcg        48
Met His Phe Ser Leu Ser Phe Ala Thr Leu Ala Leu Leu Val Ala Ser
1               5               10              15

gct gtt ggt                                                             57
Ala Val Gly


<210>  16

<211>  19

<212>  PRT

<213>  Agaricus bisporus


<400>  16

Met His Phe Ser Leu Ser Phe Ala Thr Leu Ala Leu Leu Val Ala Ser
1               5               10              15 .

Ala Val Gly


<210>  17

<211>  45

<212>  DNA

<213>  Agaricus bisporus


<400>  17
gtgggattat ggttagccaa atggtcgtag ctaattatct tgcag                      45


<210>  18

<211>  47

<212>  DNA

<213>  Agaricus bisporus


<400>  18
gtacgtaggc aagtggttcc caaaaagcac cagctaacac aacgtag                    47


<210>  19

<211>  48

<212>  DNA

<213>  Agaricus bisporus
```

<400> 19
gtacgttgaa tcgtacaaga aagtgtaatc atcctgactt tctatcag   48

<210> 20

<211> 62

<212> DNA

<213> Agaricus bisporus

<400> 20
gtaagcaccg gctgcgttcg cacacccgtc ttgtgaaaag tcgtctcatg aatatcgccc   60

ag   62

<210> 21

<211> 51

<212> DNA

<213> Agaricus bisporus

<400> 21
gtaattaacc tcatcattat tgatcctttc catgcttaca gctgttatca g   51

<210> 22

<211> 52

<212> DNA

<213> Agaricus bisporus

<400> 22
gtgcgttctt ctcttgtcct atcacaaatt ctgacgccgc aggacttgcc ag   52

<210> 23

<211> 63

<212> DNA

<213> Agaricus bisporus

<400> 23
gcaagcttct cttgtcacaa tgttaacggc gaggggtctg actcccttgg ttgtttttgt   60

tag   63

<210> 24

```
<211>   48

<212>   DNA

<213>   Agaricus bisporus


<400>   24
gtacgctata gcttgcaagg atggacatat ctaatcgggg acgtgtag              48


<210>   25

<211>   61

<212>   DNA

<213>   Agaricus bisporus


<400>   25
gtgggttcca tcatctcttg agttatgccg cggctgactg atcatgtttc taatacttca  60

g                                                                 61


<210>   26

<211>   46

<212>   DNA

<213>   Agaricus bisporus


<400>   26
gtgcgtcatg tccgtatcat ctactctcat actaatgcgc atatag                46


<210>   27

<211>   53

<212>   DNA

<213>   Agaricus bisporus


<400>   27
gtaggtccct tttatttttg gtccgattgc gtcattcatg tctatatatg cag         53


<210>   28

<211>   48

<212>   DNA

<213>   Agaricus bisporus


<400>   28
gtaggtttta ttcatccctc tgttcttatc agcttgacat ccttcaag             48
```

<210> 29

<211> 53

<212> DNA

<213> Agaricus bisporus

<400> 29
gtaagcataa ggagttgctg caggcgaccc atagcgttct aattagctgc tag   53

<210> 30

<211> 52

<212> DNA

<213> Agaricus bisporus

<400> 30
gtatgttgga actcaccgat gcgctctttg ttgattttat ttttctacat ag   52

<210> 31

<211> 50

<212> DNA

<213> Agaricus bisporus

<400> 31
gtgggtgatc attcgaggtt gtcttcctgt gtattgataa ggtttgctag   50

<210> 32

<211> 45

<212> DNA

<213> Agaricus bisporus

<400> 32
gtcagtacta gtttgttttc tcttacacct tctcattctt tgcag   45

<210> 33

<211> 56

<212> DNA

<213> Agaricus bisporus

```
<400> 33
gtgagcactg caatatggta tagcttggaa agcctttatt tatacagaac atccag          56


<210> 34

<211> 55

<212> DNA

<213> Agaricus bisporus


<400> 34
gtgagtgaag atagttctca tgtgagatcc ttgtactaat tgccgaacgt cgtag           55


<210> 35

<211> 690

<212> DNA

<213> Agaricus bisporus


<400> 35
tatatagtct ccaaattcta ttgtaatgcc attttcccaa ttcaaaagga cccgctctcg       60

aaccgggtca gatgcaattt tggtcagcaa tggtttatgt tgtttccccg taagtatgca      120

ctagaagaga acaaaacgtc actatttgct caatgcagga tgcacctggc gagataatat      180

tcttgcggtg aagtcgaaca acgtctgtag tcctgtaaaa atatacagtg agtagaggga      240

tgatgccgat gtggaaggaa gcaaccgatt acgattcgga tttggccatg agacggccgc      300

tcttgaacag accaacaata tccctttaaa tttaatacag aattactcaa tatgcttcca      360

agtatttcga ttcctcgaat tccgtctggc cgcggcatga gcatggacag gcggacagaa      420

gaggctatcg ttgtattgct tcatcagcga ccctgactag tgacttcagg catgatcatg      480

cgcttagcaa tctgtccctt caagtcgagt ccccgaattc aacagcttca acaagtcgtg      540

attatttgac ccccgactgg aatcaaattg gctcttcaaa tttcaaactt caatgcttca      600

tgcttcatgc gtcatgacgc aagctgtcaa ttttcatttt ccagttcggt cccattctca      660

ctctcgcccc tcctaatgtc ttccagaaaa                                       690


<210> 36

<211> 910

<212> DNA

<213> Agaricus bisporus


<400> 36
tatcccattt caatccctta cacgaaatct gtacttgtag tcttagaaga aagtcatgac       60

gttatataga tcacttacat agatcttcag gttttcgtag atcgacgacc gacgctctta      120
```

```
aatatttatt tacagttttt ctgttttttg ttttattgtc gcttggatat aaggtggtat      180

actttgatat gattgcctac acacatatat caacacagtt ttagttatat caacatcaaa      240

acatcagtca aggaaaacaa agagcgaacg ataaacatca gcacaagtat gtcagattat      300

ggtccaagaa cgcgaaaaga agttcgcaaa agaacagaac actatcgaaa agtgcagata      360

cataggtcac acaattaacg acttcccgga atagttccct ccaacctctt atcgcgacta      420

ctagcaccaa cggtaacacc aaaagtacct tcaggcctcc tccatccctg tgcattcaca      480

tcccaaatac tcaaatcata cctcgacaag gtcattttta cattcctagt ctctccaggg      540

ccaatcggta cagagtcgaa accgcgtagc acggaaggag gttctccagc agattcaggg      600

aagttaatgt agagttgggg agactcggca cccaaaagtc gaccggtatt cttgacgttg      660

aaggaaacct cgtacaaagg acgatggagc ctggaaaccg gtcagcgtga aattgttgat      720

tagagaaccc tcggcgctta ccaagaagcg atagaaccac cttcaacttg cgggcttgca      780

tgtccttgat cccattccct agccgcatca acttctctac ctttgacgtg cactttctta      840

atagatatct tcgagtaatc aaacttcgtg taactcaagc caaaaccaaa ttcaaaccga      900

ggctcaatgc                                                            910
```

<210> 37

<211> 800

<212> DNA

<213> Agaricus bisporus

<220>

<221> misc_feature

<222> (5)..(9)

<223> Restriction site for BglII

<220>

<221> misc_feature

<222> (790)..(795)

<223> Restriction site for KpnI

<220>

<221> misc_feature

<222> (290)..(290)

<223> n is an unknown nucleotide

```
<220>

<221>  misc_feature

<222>  (498)..(498)

<223>  n is an unknown nucleotide


<220>

<221>  misc_feature

<222>  (589)..(589)

<223>  n is an unknown nucleotide


<400>  37
gaatagatct gatgccgacc gcgggatcca cttaacgtta ctgaaatcat caaacagctt       60

gacgaatctg gatataagat cgttggtgtc gatgtcagct ccggagttga gacaaatggt      120

gttcaggatc tcgataagat acgttcattt gtccaagcag caaagagtgc cttctagtga      180

tttaatagct ccatgtcaac aagaataaaa cgcgttttcg ggtttacctc ttccagatac      240

agctcatctg caatgcatta atgcattgac tgcaacctag taacgccttn caggctccgg      300

cgaagagaag aatagcttag cagagctatt ttcattttcg ggagacgaga tcaagcagat      360

caacggtcgt caagagacct acgagactga ggaatccgct cttggctcca cgcgactata      420

tatttgtctc taattgtact ttgacatgct cctcttcttt actctgatag cttgactatg      480

aaaattccgt caccagcncc tgggttcgca aagataattg catgtttctt ccttgaactc      540

tcaagcctac aggacacaca ttcatcgtag gtataaacct cgaaatcant tcctactaag      600

atggtataca atagtaacca tgcatggttg cctagtgaat gctccgtaac acccaatacg      660

ccggccgaaa cttttttaca actctcctat gagtcgttta cccagaatgc acaggtacac      720

ttgtttagag gtaatccttc tttctagaag tcctcgtgta ctgtgtaagc gcccactcca      780

catctccacg gtacctgcag                                                  800
```